ENEE 457: Computer Systems Security
09/07/16

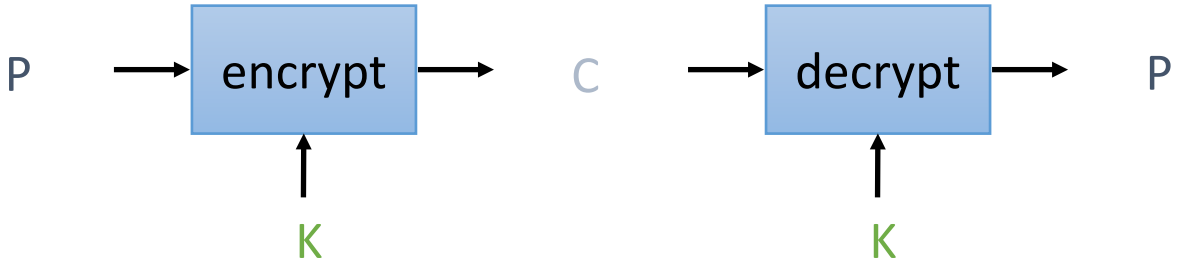
Lecture 3
Symmetric Key Encryption I:
One-Time Pad and Pseudorandom Permutations

Charalampos (Babis) Papamanthou

Department of Electrical and Computer Engineering
University of Maryland, College Park

Symmetric Cryptosystem

- Scenario
  - Alice wants to send a message (plaintext P) to Bob
  - The communication channel is insecure and can be eavesdropped
  - If Alice and Bob have previously agreed on a symmetric encryption scheme and a secret key K, the message can be sent encrypted (ciphertext C)
- Issues
  - What is a good symmetric encryption scheme?
  - What is the complexity of encrypting/decrypting?
  - What is the size of the ciphertext, relative to the plaintext?



Basic Notions

- Notation
  - Secret key K
  - Encryption function $E_K(P)$
  - Decryption function $D_K(C)$
  - Plaintext length typically the same as ciphertext length
  - Encryption and decryption are permutation functions (bijections) on the set of all n-bit arrays
- Efficiency
  - functions E_K and D_K should have efficient algorithms
- Consistency
  - Decrypting the ciphertext yields the plaintext
  - $D_K(E_K(P)) = P$

Attack on all schemes: Brute-Force Attack

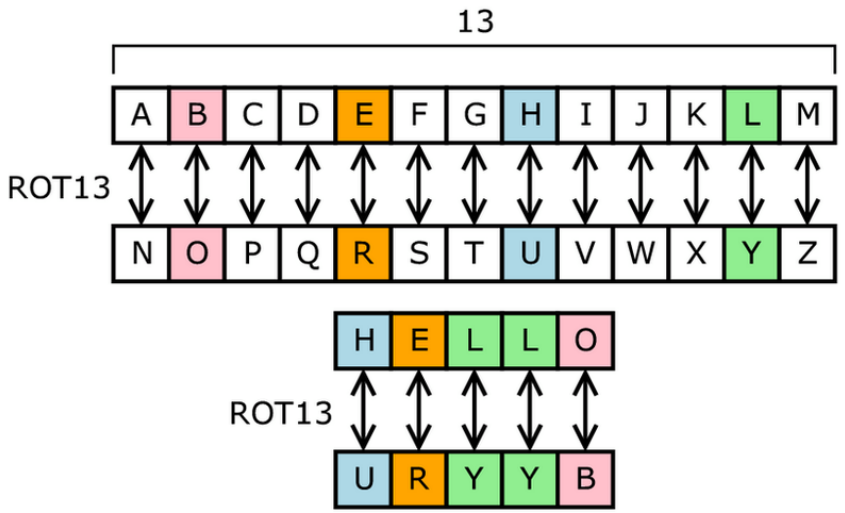
- Try all possible keys K and determine if $D_K(C)$ is a likely plaintext
  - Requires some knowledge of the structure of the plaintext (e.g., PDF file or email message)
- Key should be a sufficiently long random value to make exhaustive search attacks unfeasible

Image by Michael Cote from http://commons.wikimedia.org/wiki/File:Bingo_cards.jpg

Candidate scheme: Substitution Ciphers

- Each letter is uniquely replaced by another
- There are 26! possible substitution ciphers

- One popular substitution "cipher" for some Internet posts is ROT13



Public domain image from http://en.wikipedia.org/wiki/File:ROT13.png

Or…Substitution Boxes

- Substitution can also be done on binary numbers.
- Such substitutions are usually described by substitution boxes, or S-boxes.

	00	01	10	11
00	0011	0100	1111	0001
01	1010	0110	0101	1011
10	1110	1101	0100	0010
11	0111	0000	1001	1100

(a)

	0	1	2	3
0	3	8	15	1
1	10	6	5	11
2	14	13	4	2
3	7	0	9	12

(b)

Figure 8.3: A 4-bit S-box (a) An S-box in binary. (b) The same S-box in decimal.

Attack on Substitution ciphers: Frequency Analysis

- Letters in a natural language, like English, are not uniformly distributed
- Knowledge of letter frequencies, including pairs and triples can be used in cryptologic attacks against substitution ciphers

a:	8.05%	b:	1.67%	c:	2.23%	d:	5.10%
e:	12.22%	f:	2.14%	g:	2.30%	h:	6.62%
i:	6.28%	j:	0.19%	k:	0.95%	l:	4.08%
m:	2.33%	n:	6.95%	o:	7.63%	p:	1.66%
q:	0.06%	r:	5.29%	s:	6.02%	t:	9.67%
u:	2.92%	v:	0.82%	w:	2.60%	x:	0.11%
y:	2.04%	z:	0.06%				

Letter frequencies in the book *The Adventures of Tom Sawyer*, by Twain.

What would a great symmetric encryption scheme satisfy?

- What if we could devise a system such that we can encrypt and the ciphertext does not reveal anything about the plaintext (apart from its length)
- Let’s express it mathematically

Perfect security

- Pick messages m_1 and m_2
- Pick a ciphertext c
- Encrypt m_1
- Encrypt m_2
- Compute the probability Pr[Enc(m_1)=c] (over the choice of the random key)
- Compute the probability Pr[Enc(m_2)=c] (over the choice of the random key)
- Enc is secure if for all messages m_1 and m_2 and for all ciphertexts c
  - Pr[Enc(m_1)=c]= Pr[Enc(m_2)=c]

One-Time Pads: Achieving perfect security

- There is one type of substitution cipher that is absolutely unbreakable
  - The **one-time pad** was invented in 1917 by Joseph Mauborgne and Gilbert Vernam
  - We use a block of shift keys, $(k_1, k_2, \ldots, k_n)$, to encrypt a plaintext, M, of length n, with each shift key being chosen uniformly at random
- Since each shift is random, every ciphertext is equally likely for any plaintext

Algorithms of one-time pad

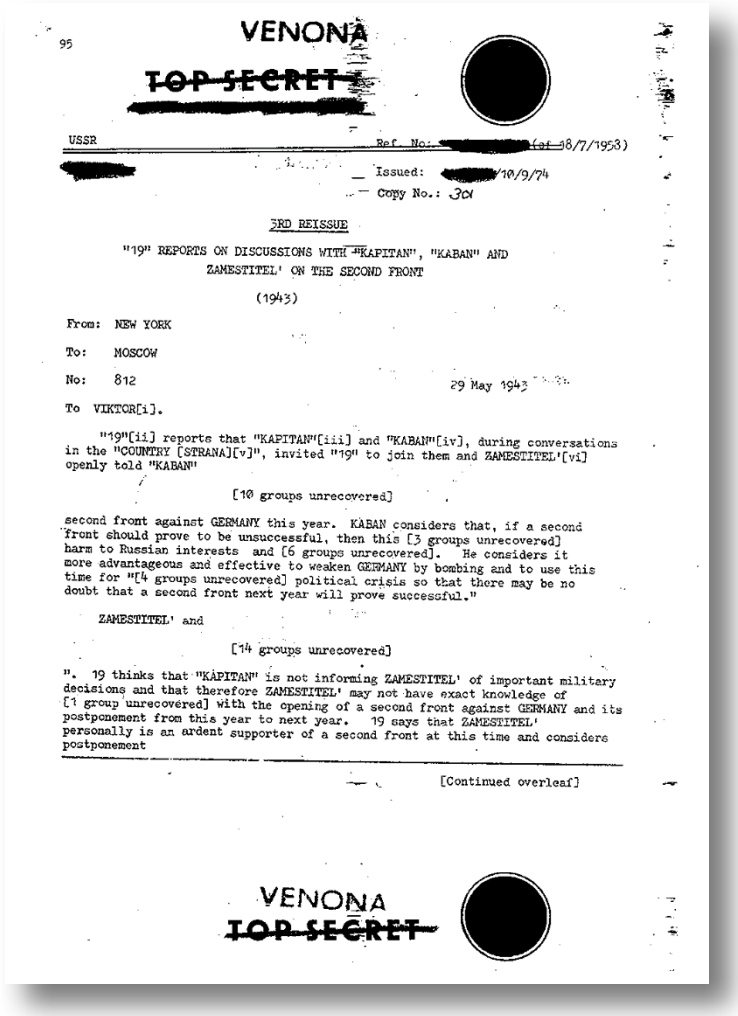
- $K \leftarrow KeyGen(n)$: Pick a random key K of n bits
- $E_K(A)$: On input plaintext A, compute ciphertext B=A XOR K
- $D_K(B)$: On input ciphertext B, compute plaintext A=B XOR K
- Correctness: B XOR K= (A XOR K) XOR K= A XOR 0 = A
- Security?
  - Note that $Enc_K(m_1)=c$ is the event m_1 XOR K = c which is the event K = m_1 XOR c
  - K is chosen at random (irrespective of m_1 and m_2, and therefore the probability is 2^{-n}
  - Namely ciphertext does not reveal anything about the plaintext

Key space and message space in one-time pad

- Key space should be at least equal to the message space
- Suppose not and the key space is missing one element and does not contain 0000000…00
- For a given c, there exists a message m such that
  - Pr[Enc(m) = c]=0
  - E.g., If key does not contain 00000000000…00, then m = c
  - But for all other messages m' that are not equal to c we have that
    - Pr[Enc(m') = c]>0=1/(2^{n}-1) (why is that?)
  - Therefore the definition does not hold.
  - In particular, if I see a ciphertext, I have excluded one possibility

One-time pad is not practical

- In spite of their perfect security, one-time pads have some weaknesses
- The key has to be as long as the plaintext
- Keys can never be reused
  - Repeated use of one-time pads compromised communications during the cold war



95

VENONA

~~TOP SECRET~~

USSR

Ref No: [redacted] (of 18/7/1953)

[redacted]

Issued: [redacted] 10/9/74

Copy No.: 30

3RD REISSUE

"19" REPORTS ON DISCUSSIONS WITH "KAPITAN", "KABAN" AND ZAMESTITEL' ON THE SECOND FRONT

(1943)

From: NEW YORK

To: MOSCOW

No: 812

29 May 1943

To VIKTOR[i].

"19"[ii] reports that "KAPITAN"[iii] and "KABAN"[iv], during conversations in the "COUNTRY [STRANA][v]", invited "19" to join them and ZAMESTITEL'[vi] openly told "KABAN"

[10 groups unrecovered]

second front against GERMANY this year. KABAN considers that, if a second front should prove to be unsuccessful, then this [3 groups unrecovered] harm to Russian interests and [6 groups unrecovered]. He considers it more advantageous and effective to weaken GERMANY by bombing and to use this time for "[4 groups unrecovered] political crisis so that there may be no doubt that a second front next year will prove successful."

ZAMESTITEL' and

[14 groups unrecovered]

". 19 thinks that "KAPITAN" is not informing ZAMESTITEL' of important military decisions and that therefore ZAMESTITEL' may not have exact knowledge of [1 group unrecovered] with the opening of a second front against GERMANY and its postponement from this year to next year. 19 says that ZAMESTITEL' personally is an ardent supporter of a second front at this time and considers postponement

[Continued overleaf]

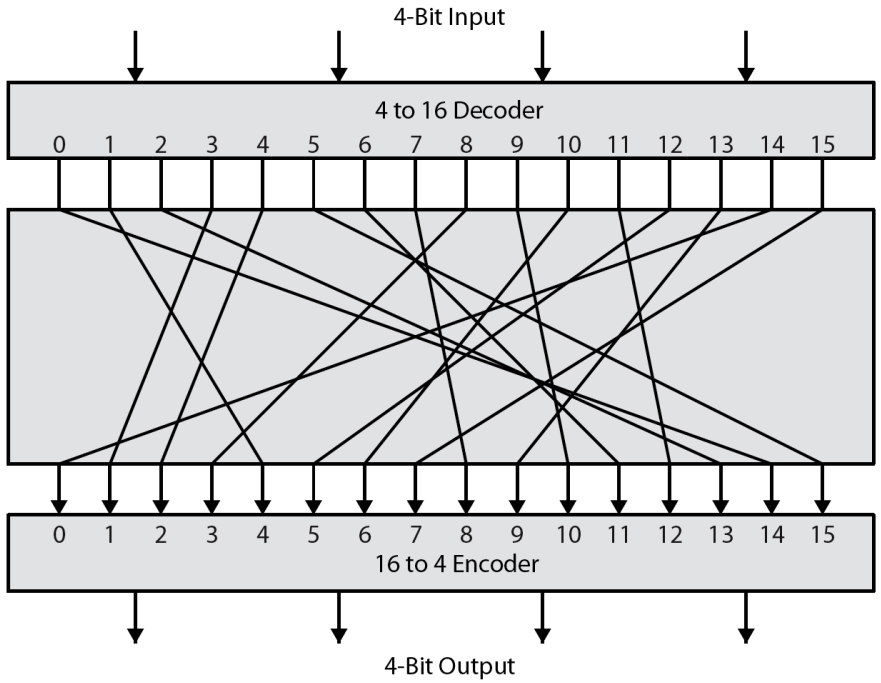
VENONA

~~TOP SECRET~~

Public domain declassified government image from https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/venona-soviet-espionage-and-the-american-response-1939-1957/part2.htm

What do we want to use in practice?

- Usability: Same key all the time



- Is this enough? What are two problems?

What do we want to use in practice?

- Size: Small key (128 bits)
- Security: It turns out that **perfect secrecy** is very strong if we want to achieve both small key and one key
  - How about if we improve the best strategy of the attacker, which is still going to be really bad for practical purposes

- Answer: **Computational Secrecy**
- **Intution: The ciphertext does not reveal anything about the plaintext as long as our attacker runs in time polynomial (like all machines in this class)**
- If attacker can run in time 2^{size_of_key}, all bets are off
- But this is too long…

Pseudorandom permutations (PRP)

- P_k(x) = y, where x\in X, y\in Y and k\in K.
- A bijection (therefore domain and range are the same)
- k not necessarily from the same domain
- E.g., k can be 128 bits, while x and y can be 256 bits
- **Output is pseudorandom** (cannot be distinguished from a truly random number)
- Assume we have a P_k(x). How do we build a semantically-secure symmetric encryption scheme?
  - Pick a random r
  - Enc_k(x) = P_k(r) XOR x , r
  - Dec_k(c) = P_k(r) XOR c
- How can you prove the above is semantically-secure?
  - Prove that if attacker can distinguish, then he can break the PRP